Football Club Transfer Networks and Performance

Dennis Coates (UMBC and NRU HSE)

Iuliia Naidenova (NRU HSE)

Petr Parshakov (NRU HSE)

Introduction

Nowadays football becomes a global business that attracts billions of euros every year. One of the main resources in football is talented players. Professional football clubs compete for the best players that are a key source of clubs' success in national and international competitions (Liu et al., 2016; Rossetti and Caproni, 2016). Moreover, better performance and star players attract supporters to the club both in a domestic country and abroad. This, in turn, leads to the higher financial performance of the club in the form of revenues and sponsorship contracts (Naidenova et al., 2016).

Globalization has upended football labor market. Nowadays football clubs can use different strategies on players' labor market: buying famous players through transfer markets or scouting and growing talented youths (Lee et al., 2015). The aspect of selling players is also important. Some clubs sell players to buy better ones, whereas other clubs are focused on selling uncovered young talents to earn money (Pannenborg, 2010; Rossetti and Caproni, 2016). Therefore, the quality of transfer management affects both sports and financial performance of the club.

There are different dimensions of transfer (or talent) strategy of football clubs. First, club management is deciding on buying vs. growing new players. Second, the club is also deciding on buying players (both mature and young) abroad or in a domestic league. The first dimension can be captured by team average age. The effect of age as an indicator of human capital for the football team performance is investigated by Dawson et al. (2000). The second dimension is analyzed in the studies of diversity impact on team performance, which has been the focus of several papers

(Brandes et al., 2009; Franck and Nuesch, 2010; Ingersoll et al., 2014; Kahane et al., 2013; Frick and Rose, 2017). However, age and diversity are indirect indicators of football club transfer strategy: low team age or low doesn't necessary imply that the club hasn't bought the players. In this study, we use a network approach to analyze how characteristics of football club's player transfer network influence on club's performance. European transfer market is considered as a weighted network. We evaluate different measures of centrality as indicators of football club strategy in transfers to analyze its impact on sports and financial performance (financial performance is studied only for the English Premier League). We focus on European football for three reasons. First, it is the world's most popular sport (Matheson, 2003). Second, the frequency of transfers is high relative to the other sports. Third, the distribution of teams participating in the market is broad (Matheson, 2003).

Previous research on networks analysis in football found mixed results: positive correlation between match performance and betweenness and closeness centrality of transfer network (Liu et al., 2016) and possible non-linear relationship with roster changes (Rossetti and Caproni, 2016). The influence of transfer network on club's financial performance is poorly investigated. The results of the study (Liu et al., 2016) show the weak relation between club's revenues and transfer network features. However, these authors use correlation analysis and do not control for other determinants of performance of football club like the quality of players. In our paper, we use the massive dataset on transfers and club characteristics to study the causal relation between transfer strategy and performance.

The paper begins with a description of networks analysis and related indicators of European football clubs. In next section, we analyze the relation between sports performance of football club. Implications for financial performance are studied on the English Premier Leagues (EPL) clubs due to the availability of financial reports. The paper ends with a presentation and discussion of empirical results, robustness checks, and a conclusion.

Football club networks and centrality metrics

The number of papers which use network analysis and its application has grown since the major methodological breakthroughs in the 1980s. Wasserman and Galaskiewicz (1994) provide an extensive review of networks analysis application for social sciences.

In our study, we use data on transfers between European football clubs as a weighted directed network (graph). If there is a transfer from club A to club B, we consider them connected with the direction from A to B. The number of transfers from club A to club B is considered as the weight of the connection.

To build such networks we use data for the Belgium Jupiler League, England Premier League, France Ligue 1, Germany Bundesliga, Italy Serie A, Netherlands Eredivisie, Poland Ekstraklasa, Portugal Liga ZON Sagres, Scotland Premier League, Spain LIGA BBVA and Switzerland Super League for the seasons 2008 to 2016. We use dataset provided by Hugo Mathien (Mathien, 2016). It should be noted, that the transfer data is constructed from the match-level data on line-ups. So, if player A participates at least one match playing for team one in a particular season, and in next season he participates at least one match for another team, we consider this as a transfer. With such an approach, we do not distinguish between loans and transfers, which is the limitation of our study. We address this limitation in the robustness checks section.

Figure 1 represents the transfer network of the clubs. Portuguese, Spain and English leagues are in the center of the network and they seem to be widely connected with the other leagues. On the contrary, there are leagues with a low number of connections with other leagues. Interestingly, there are both top- and low-level leagues among them: for example, Italian and Polish leagues.

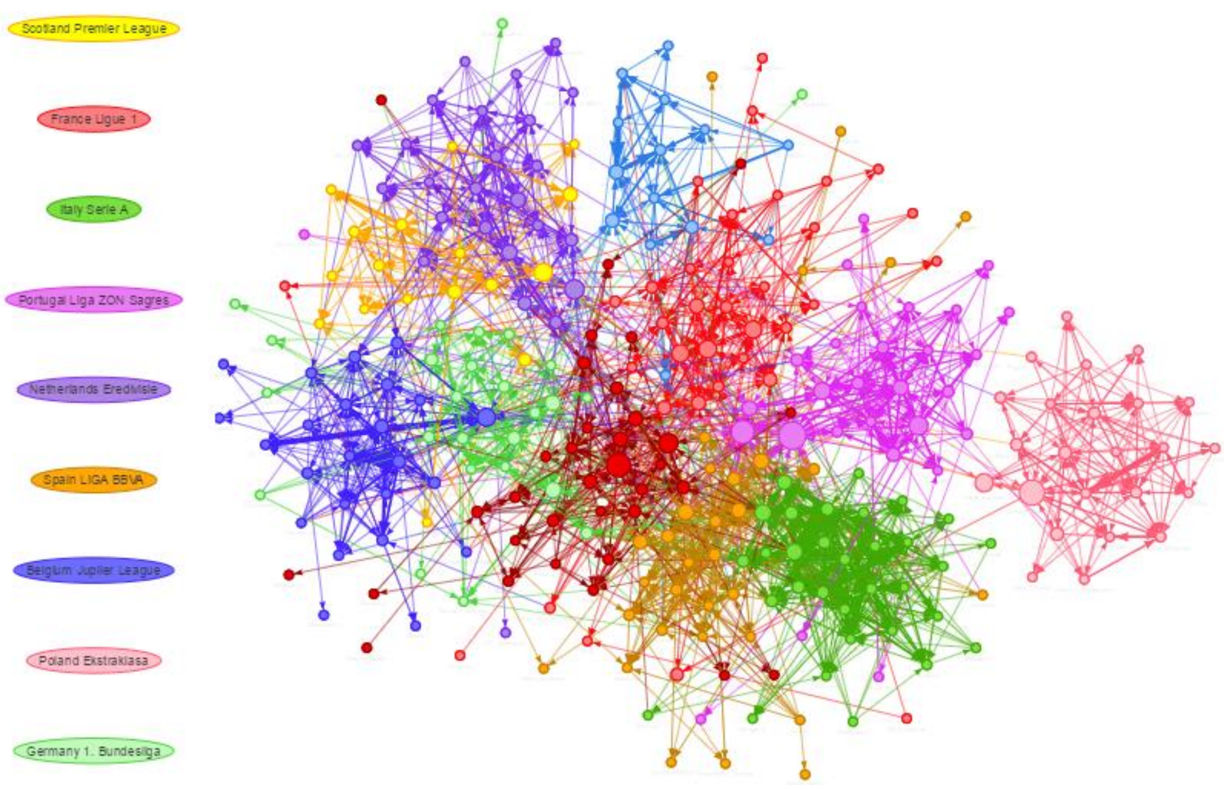

Figure 1. Transfer network of the clubs. The color represents the league.

To quantify the transfer strategy of the clubs, we use three network indicators. The degree of a vertex is its most basic structural property, the number of its adjacent edges. We evaluate both out-degree for the paths from a vertex, in-degree for measures paths to a vertex and total degree for the sum of the two (Freeman, 1978). Betweenness centrality is defined by the number of shortest paths going through a vertex or an edge. We use an algorithm proposed by Brandes (2001) in order to calculate it. We also evaluate closeness centrality, which measures how many steps is required to access every other vertex from a given vertex (Freeman, 1978). Formally, the closeness centrality of a vertex is defined by the inverse of the average length of the shortest paths to and from all the other vertices in the graph. However, we do not use it for further analysis despite it is popular in network application studies, since it might be correctly interpreted only for a connected network, which is not our case. All network measures we use are normalized in order to make them comparable across different types of networks. We suppose all three network characteristics play an important role for sports performance of the club because it reflects the quality of scout system. However, since building scout system might be costly, it might negatively affect the financial performance.

In order to analyze an evolution of transfer network, we construct all metrics using rolling window. The size of a window is three years. So, in order to get centrality measures for the season 2010/2011, we use data on transfers for the season 2008/2009, 2009/2010 and 2010/2011. After that, we move one step ahead and use data for the next three seasons. As a results, we have panel data on networks characteristic, which varies from season to season.

We analyze three types of networks among the same clubs. First, we evaluate centrality measures for the networks which consist of all transfers. Next, we use data only for transfers between leagues to build the network. The idea is to understand the embeddedness of a club to the global labor market of players. Finally, we use data only on transfers between clubs of the same league to measure understand the orientation on the local market. Table 1 represents basic descriptive statistics for the network measures. Three panels of the table refer to three networks described above. All centrality measures are normalized in order to compare metrics for different types of networks.

As one can see from the table, the minimum value for in- and out-degree for all panels is zero. So, we observe clubs which only sell or buy players during the three-year window. The percentage of such clubs is 1% and 8%, respectively. Interestingly, degree centrality varies a lot among the types of the network while betweenness is more robust. That might indicate that betweenness is a more reliable indicator of football club strategy: those clubs who choose to be "mediators" are doing it both on the global and local labor market. Comparing in- and out-degree of the panel (b) and (c) one can conclude that there are more transfers inside the same league.

Table 1. Descriptive statistics of network characteristics

Variable	Obs	Mean	Std. Dev.	Min	Max
			(a) All transfers		
degree	1,106	0.072	0.034	0.004	0.225
in-degree	1,106	0.039	0.019	0	0.145
out-degree	1,106	0.033	0.021	0	0.123
betweenness	1,106	0.010	0.009	0	0.071
			(b) Transfers between leagues		
degree	1,031	0.032	0.024	0.005	0.163
in-degree	1,031	0.017	0.014	0	0.084
out-degree	1,031	0.015	0.014	0	0.116
betweenness	1,031	0.010	0.014	0	0.097
			(c) Transfers inside leagues		
degree	1,105	0.044	0.022	0.004	0.141
in-degree	1,105	0.024	0.013	0	0.083
out-degree	1,105	0.020	0.013	0	0.066
betweenness	1,105	0.000	0.000	0	0.003

In order to illustrate the importance of network measures as indicators of transfer strategy, we perform k-mean cluster analysis for the clubs of EPL as one of the most popular leagues. We concentrate here only on one league for the purpose of visualization. Figure 2 shows the results. The left graph shows clusters according to points per game as an indicator of sports performance and revenue as an indicator of financial performance. We report descriptive statistics on these indicators later in Table 2. The right graph shows four clusters according to points per game, revenue and network characteristics. The number of clusters was chosen according to the explained share of variance.

The most interesting aspect of this Figure is that the number and the composition of clusters differ while taking network measures into account. This indicates that these metrics contains information which is not represented in the sports and financial indicators. In other word, clubs with the same sports and financial performance might have totally different transfer structure, which should be taken into account while analyzing, for example, the efficiency of team management. For example, on the left graph of Figure 2 Liverpool and Chelsea are in the same cluster, while these clubs have totally different strategies in terms of players. Liverpool owner, John Henry, is the owner of Boston Red Sox and is a well-known supporter of Moneyball approach to team management. On the right graph, Liverpool is in another cluster, which contains clubs, known as “mediators”, which are buying underestimated players and selling them for the moneybags.

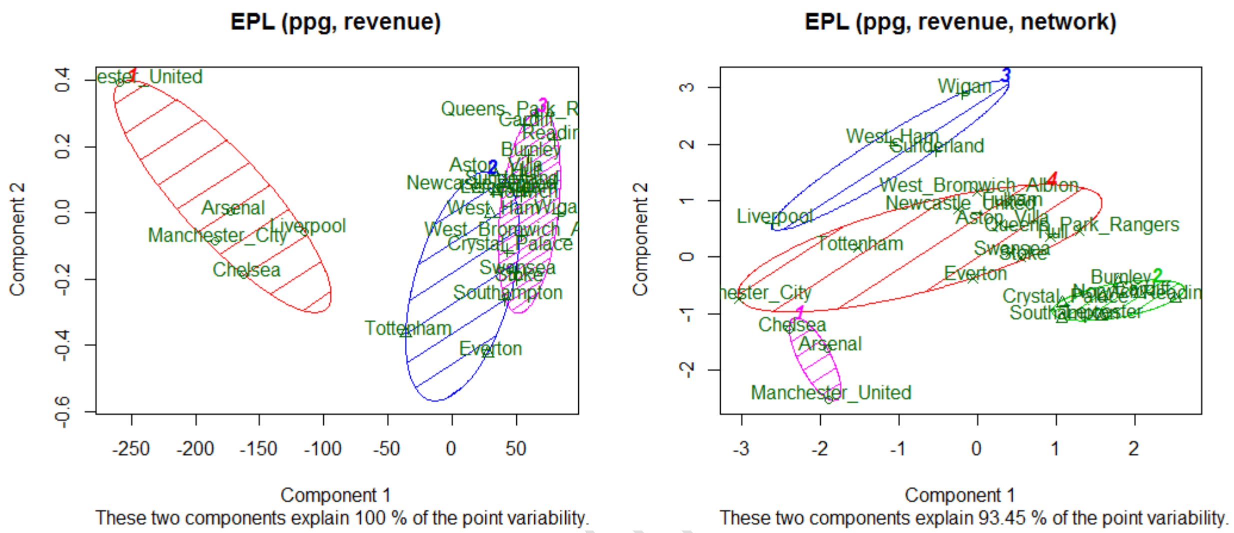

Figure 2. Clusters of EPL clubs.

Team performance and transfer strategy

Methodology and data

To analyze the impact of team performance to transfer strategy we use average point per game as an indicator of sports performance. Most commonly, researches use win percentage (Espitia-Escuer and GarcÍa-CebriÁn, 2004), however, since draws are frequent in football, this indicator might give a biased impression of performance (Dawson et al., 2000). Following the study of Hausman and Leonard (1997), we use football club revenue as an indicator of financial performance, because profit indicators are affected by several unobservable factors[1].

Since it is important to control for the determinants of team performance in order to capture the effect of transfer strategy, we use a set of control variables. First, we measure players quality by estimation of their skill from FIFA video game simulator, developed by EA. Sherif (2016)

[1] Results for the profit as a dependent variable are available upon a request.

explains the process of evaluating the skills. It is the multistage process, on the first step a "network of over 9000 members reviews the player's abilities, watch him play, and help assign him various ratings." On the next stage, this data is then reviewed "by 300 editors, which arrange it into 300 fields and 35 attribute categories." After that EA "uses this feedback in conjunction with its own stats (scoured from other agencies) to determine ratings." We can consider this rating as a result of a massive survey with a number of respondents. The final rating is distributed from 1 to 100.

Another important driver of team performance is coach quality. Together with team management coach decides on transfers, he motivates players and chooses proper tactics for a particular game and training strategy. So, coach influence team performance (Dawson et al., 2000; Hentschel et al., 2012; Paola and Scoppa, 2012). In some studies, the authors elaborate a metric to measure coach quality (Kahn, 1993; Pfeffer and Davis-Blake, 1986). Another approach is to use fixed effect since coach turnover is low (Borland and Lye, 1996; Dawson et al., 2000). In our paper, we control for the coach effect by accounting the variation in tactics. For each game, we have coordinates of players of the starting squad. We use Y-axis coordinate of each player and average it by a team. Next, we find the standard deviation of this averages across the matched during the season. The idea is to capture the variability of tactics during the season, which partly represents the coach responsibilities. We also include team fixed effects in the regression to control for the other sources of unobservable heterogeneity.

Table 2. Descriptive statistics of football club characteristics

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Rating	1,106	70.001	4.927	57.143	90
Strategy variation	1,106	2.823	1.556	0	7.240
Points per game	1,106	1.343	0.431	0.118	2.684
Revenue (mln GBR)	60	153.217	102.530	56	433
Attendance	60	36,211.35	14,024.97	17,779	75,53

Table 2 contains descriptive statistics of these indicators. The variation in rating is high due to the variation in clubs and leagues. Average points per game are 1.3 indicating the fact that our sample includes significant number of draws.

To analyze the impact of transfer strategy on the team performance we estimate the following regression equations:

$$points\_per\_game_{it} = \beta_0 + \beta_1 \cdot \mathbf{NETW}_{\boldsymbol{it}} + \beta_2 \cdot \mathbf{CV_{it}} + \phi_i + \epsilon_{it} \qquad (1)$$

$$revenue_{it} = \beta_0 + \beta_1 \cdot \mathbf{NETW}_{\boldsymbol{it}} + \beta_2 \cdot \mathbf{CV_{it}} + \phi_i + \epsilon_{it}. \qquad (2)$$

Where **NETW** is a vector of networks metrics described above, **CV** is a vector of control variables consists of the average player and variation in strategy, ϕ is a team fixed effect and ϵ is the error term. Note that for the second equation average season attendance is added to the set of control.

Empirical results

Tables 3 and 4 represents the results of the regression analysis. In Table 3 we report the results for the model with all transfers, in Table 4 we consider the type of transfer: between leagues or inside the same league. The second model in both tables contains the results for the EPL. Note that the number of observation for these models is much lower.

The coefficients for the player rating are as one would expect. It positively affects points per game and does not affect the revenue. Surprisingly, the coefficient of the strategy variation is negative. It seems that changing tactics a lot is harmful to the team performance. However, such result might also indicate that the best teams do not need to change the tactics from match to match.

Network measures, which are at the core of this study, are jointly significant in all four regression models. Summarizing the statistical significance and marginal effects of them, we formulate three main results of this study. First, it's better to hire players from a limited number of clubs: all three in-degree metrics are statistically significant and negative. The marginal effect of in-degree between leagues and inside the same league are close, so the is no benefit of buying players from the same league. Second, being an important intermediary negatively affects sports results. Probably, such clubs often sell their best players to rich clubs. Third, teams financially benefit from being a global intermediator and suffer from being local intermediator. So, taking all the results into consideration, we can formulate recommendations for the club who is maximizing financial performance. The club should have low in-degree between leagues and in the same league, high betweenness between leagues, high out-degree in the same league and low betweenness in the same league. Figure 3 contains visualization of optimal transfer strategy: red circle represents the club with an optimal strategy. So, the club should minimize risks by finding a partner in the other league and sell the players to the different clubs in the same league.

Table 3. Regression results for the network measures for all transfers

	(1)	(2)
League	All	EPL
Dependent variable	Point per game	Revenue
In-degree	-2.2512***	156.3558
	(0.762)	(319.456)
Out-degree	0.6165	1,066.0922**
	(0.632)	(404.114)
Betweeness	-2.5548*	-1,333.9208
	(1.538)	(1,103.047)
Average FIFA rating	0.0383***	-1.4231
	(0.005)	(4.129)
Strategy variation	-0.0196***	-3.1350
	(0.007)	(3.420)
Constant	-1.1886***	240.5152
	(0.385)	(306.679)
Observations	1,106	60
R-squared	0.107	0.204
Number of teamid	272	25

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Regression results for the network measures for the transfers between leagues and inside the same league

	(1)	(2)
League	All	EPL
Dependent variable	Point per game	Revenue
In-degree between leagues	-2.2087**	-706.0259**
	(0.997)	(282.211)
Out-degree between leagues	0.6660	-17.0740
	(1.125)	(249.957)
Betweeness between leagues	-0.6826	551.3632**
	(1.306)	(237.006)
In-degree in same league	-2.7172**	-569.7252*
	(1.229)	(278.553)
Out-degree in same league	0.5637	623.9566*
	(1.009)	(345.276)
Betweenness in the same league	-37.0140	-11,150.6490*
	(31.803)	(6,128.056)
Average FIFA rating	0.0403***	0.1111
	(0.006)	(1.577)
Strategy variation	-0.0204***	
	(0.007)	
Average attendance		-0.0032
		(0.003)
Season dummies	Included	Included

Constant	-1.3151***	258.7505*
	(0.428)	(129.540)
Observations	1,030	59
R-squared	0.113	0.828
Number of teamid	259	24

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

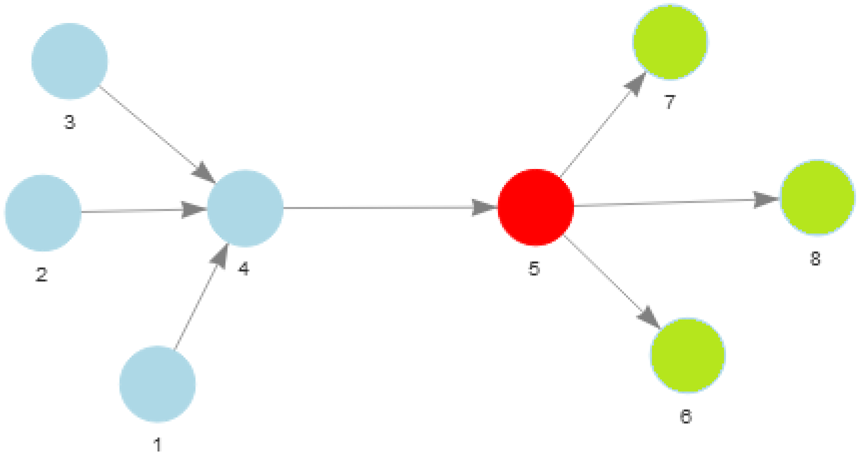

Figure 3. Visualization of optimal transfer strategy. The red circle represents a club with an optimal strategy.

A set of robustness check is performed. We use quantile regression approach to understand if our results differ for the different quantiles of clubs according to points per game. The results for the median, 0.25 and 0.75 are almost the same even in terms of marginal effects. We also interact the network characteristics with dummy indicators of top leagues and top clubs. Despite these interactions are jointly significant, our results do not change a lot. Finally, since we do not observe all transfers (for example, from South-American leagues), we test if our network metrics are stable. For this purpose, we perform a set of simulations: we randomly exclude some percentage of clubs, re-evaluate network measure and re-estimate regressions. Our results are robust in terms of the magnitude to the 7% percent of exclusion of clubs and to the 27% of the clubs in terms of the sign of all networks measures in regression equations. The results are available upon a request.

Our study is subject to at least three limitations. First, we don't observe transfers from all leagues and lower divisions. Second, we don't observe players who move to another team but do not participate matches. Third, financial data is available only for one league.

References

Borland, J., Lye, J., 1996. Matching and mobility in the market for Australian rules football coaches. ILR Rev. 50, 143–158.

Brandes, L., Franck, E., Theiler, P., 2009. THE EFFECT FROM NATIONAL DIVERSITY ON TEAM PRODUCTION - EXPIRICAL EVIDENCE FROM THE SPORTS INDUSTRY. Schmalenbach Bus. Rev. SBR 61, 225–246.

Brandes, U., 2001. A faster algorithm for betweenness centrality*. J. Math. Sociol. 25, 163–177.

Dawson, P., Dobson, S., Gerrard, B., 2000. Stochastic frontiers and the temporal structure of managerial efficiency in English soccer. J. Sports Econ. 1, 341–362.

Espitia-Escuer, M., GarcÍa-CebriÁn, L.I., 2004. Measuring the efficiency of Spanish first-division soccer teams. J. Sports Econ. 5, 329–346.

Franck, E., Nuesch, S., 2010. The Effect of Talent Disparity on Team Productivity in Soccer. J. Econ. Psychol. 31, 218–229.

Freeman, L.C., 1978. Centrality in social networks conceptual clarification. Soc. Netw. 1, 215–239.

Frick, B., Rose, A., 2017. Over the Top: Team Composition and Performance in Himalayan Expeditions.

Hausman, J.A., Leonard, G.K., 1997. Superstars in the National Basketball Association: Economic value and policy. J. Labor Econ. 15, 586–624.

Hentschel, S., Muehlheusser, G., Sliwka, D., 2012. The impact of managerial change on performance: The role of team heterogeneity.

Ingersoll, K., Malesky, E., Saiegh, S.M., 2014. Diversity and Group Performance: Evidence from the World’s Top Soccer League. Conf. Pap. -- Am. Polit. Sci. Assoc. 1–N.

Kahane, L., Longley, N., Simmons, R., 2013. The effects of coworker heterogeneity on firm-level output: assessing the impacts of cultural and language diversity in the National Hockey League. Rev. Econ. Stat. 95, 302–314.

Kahn, L.M., 1993. Managerial Quality, Team Success, and Individual Player Performance in Major League Baseball. Ind. Labor Relat. Rev. 46, 531–547. doi:10.1177/001979399304600306

Lee, S., Hong, I., Jung*, W.-S., 2015. A Network Approach to the Transfer Market of European Football Leagues. New Phys. Sae Mulli 65, 402–409. doi:10.3938/NPSM.65.402

Liu, X.F., Liu, Y.-L., Lu, X.-H., Wang, Q.-X., Wang, T.-X., 2016. The Anatomy of the Global Football Player Transfer Network: Club Functionalities versus Network Properties. PLOS ONE 11, e0156504. doi:10.1371/journal.pone.0156504

Matheson, V.A., 2003. European football: a survey of the literature. Williams College, Department of Economics.

Mathien, H., 2016. European Soccer Database [WWW Document]. Kaggle. URL https://www.kaggle.com/hugomathien/soccer (accessed 11.11.16).

Naidenova, I., Parshakov, P., Chmykhov, A., 2016. Does football sponsorship improve company performance? Eur. Sport Manag. Q. 16, 129–147. doi:10.1080/16184742.2015.1124900

Pannenborg, A., 2010. Big Men, Big Gains? The Involvement of African Club Officials in the Transfer of Players. Afr. Hist. Rev. 42, 63–90. doi:10.1080/17532523.2010.483804

Paola, M.D., Scoppa, V., 2012. The Effects of Managerial Turnover: Evidence from Coach Dismissals in Italian Soccer Teams. J. Sports Econ. 13, 152–168. doi:10.1177/1527002511402155

Pfeffer, J., Davis-Blake, A., 1986. Administrative succession and organizational performance: How administrator experience mediates the succession effect. Acad. Manage. J. 29, 72–83.

Rossetti, G., Caproni, V., 2016. Football Market Strategies: Think Locally, Trade Globally.

Sherif, S., 2016. EA explains how FIFA player ratings are calculated. VG247.com.

Wasserman, S., Galaskiewicz, J., 1994. Advances in Social Network Analysis: Research in the Social and Behavioral Sciences. SAGE Publications.